ALGEBRAIC MEASURES OF ENTANGLEMENT

JEAN-LUC BRYLINSKI

Abstract. We study the rank of a general tensor u in a tensor product space $H_1 \otimes \cdots \otimes H_k$. The rank of u is the minimal number p of pure states $v_1, \cdots, v_p$ such that u is a linear combination of the v_j's. This rank is an algebraic measure of the degree of entanglement of u. Motivated by quantum computation, we completely describe the rank of an arbitrary tensor in $(\mathbb{C}^2)^{\otimes 3}$ and give normal forms for tensor states up to local unitary transformations. We also obtain partial results for $(\mathbb{C}^2)^{\otimes 4}$; in particular, we show that the maximal rank of a tensor in $(\mathbb{C}^2)^{\otimes 4}$ is equal to 4.

1. Rank of a tensor

Let H be a complex Hilbert space; the hermitian scalar product will be denoted by $\langle u|v\rangle$. It is complex linear in v and antilinear in u. A state in a complex H is an element of the projective space $\mathbb{P}(H)$. The points of $\mathbb{P}(H)$ can be viewed alternatively as complex lines in H, or as elements of the unit sphere $S(H)$ up to the scaling action of complex numbers $e^{i\alpha}$. We will use the mathematical notation for states (u, v, etc...) as opposed to kets. The state u gives the projection operator $P_u : H \to H$ where $P_u(v) = \langle u|v\rangle$. P_u is an idempotent hermitian operator of rank 1; in this way we realize $\mathbb{P}(H)$ as the orbit of the unitary group comprised of such operators.

In quantum mechanics, the combination of several quantum systems corresponds to the Hilbert space tensor product $E = H_1 \otimes \cdots \otimes H_k$ of the relevant Hilbert spaces. A state u in E is called pure if it is a tensor product $\phi_1 \otimes \phi_2 \cdots \phi_k$ of states; otherwise it is called *entangled*. It is easy to characterize pure states in terms of homogeneous quadratic equations for the components of the tensor u. If we pick orthonormal bases of each H_j and write $u_{a_1 \cdots a_k}$ for the components of u, we have

Proposition 1.1. *The state u in $E = H_1 \otimes \cdots \otimes H_k$ is pure iff the following "exchange property" is verified: for any k-tuples $(a_1, \cdots, a_k)$, $(b_1, \cdots, b_k)$, $(c_1, \cdots, c_k)$, $(d_1, \cdots, d_k)$ such that for each j, (c_j, d_j) is a permutation of (a_j, b_j) we have*

$$u_{a_1,\cdots a_k} u_{b_1,\cdots b_k} = u_{c_1,\cdots,c_k} u_{d_1,\cdots d_k} \tag{1.1}$$

Proof. Clearly a pure tensor satisfies the exchange property. To prove the converse, we proceed by induction over k. We pick a basis $(e_0, \cdots, e_m)$ of H_1 and write $u = \sum_i e_i \otimes v_i$ where $v_i \in H_2 \otimes \cdots \otimes H_k$. If $v_i \neq 0$ for some i, the exchange property for the case $a_1 = b_1 = i$ implies that v_i satisfies the exchange property, so is a pure tensor by the inductive hypothesis. Next, if v_i and v_j are non-zero, we apply the exchange property to

1991 *Mathematics Subject Classification.* 81P68, 68Q17, 15A69, 14N15.
Key words and phrases. tensor states, rank, entanglement, determinant, hyperdeterminant.
I thank Joseph Bernstein and Ranee Brylinski for very useful discussions.
Research supported in part by NSF Grant No. DMS-9803593.

the case where $a_1 = d_1 = i$, $b_1 = c_1 = j$, $a_l = c_l$ and $b_l = d_l$ for $l \geq 2$, and conclude that the tensors v_i and v_j are proportional. It follows that u is a pure tensor. $\square$

Geometrically, the set of pure states is a closed complex algebraic subvariety of $\mathbb{P}(H)$, isomorphic to the product $\mathbb{P}(H_1) \times \cdots \times \mathbb{P}(H_k)$, which is known as the Segre product. So its dimension is $d_1 + \cdots + d_k - k$, where $d_j = \dim(H_j)$. Accordingly, the pure tensors in $E = H_1 \otimes \cdots \otimes H_k$ form a closed complex algebraic subvariety of dimension $d_1 + \cdots + d_k - k + 1$.

Entangled states occur naturally in classical algorithms for matrix multiplication [Str1] [Str2]. They are very important in quantum mechanics; cf. e.g. the famous Einstein-Podolsky-Rosen work. Quantum computation lives in the tensor product Hilbert spaces $(\mathbb{C}^2)^{\otimes n}$, and states used in quantum coding and quantum teleportation are typically quite entangled (see e.g. [C-R-S-S] [Go] [Ste]). So it seems important to study how entangled states can be. The following is a classical notion (see [B-C-S]).

Definition 1.2. We say a state u in $E = H_1 \otimes \cdots \otimes H_k$ has rank $\leq p$ if we can write

$$u = \sum_{j=1}^{p} \lambda_j v_j \tag{1.2}$$

where each v_j is a pure state.

A natural question is to find the highest rank of all states in E; we can only answer this in very special cases. At least we can give a lower bound:

Proposition 1.3. *Let H_j be vector spaces of dimension d_j. Then the highest rank of states in $E = H_1 \otimes \cdots \otimes H_k$ is at least equal to the rational number*

$$\frac{d_1 d_2 \cdots d_k}{d_1 + d_2 + \cdots + d_k - k + 1} \tag{1.3}$$

For instance, take $k = 3$, $d_1 = 3$, $d_2 = 4$, $d_3 = 5$; then the highest rank is at least $3 \times 4 \times 5/10 = 6$.

In case $k = 2$, it is easy to describe this degree of entanglement of any state in classical terms:

Proposition 1.4. *The degree of entanglement (rank) of a state u in $E = H_1 \otimes H_2$ is the rank of the matrix u_{ab}.*

In particular, for $k = 2$, the degree of entanglement gives a nice stratification of projective space $\mathbb{P}(E)$. Let S_p denote the set of states of rank $\leq p$. Then S_p is a closed algebraic subvariety of $\mathbb{P}(E)$, defined as the vanishing locus of all order $p+1$ minors of the matrix u_{ab}. The singular locus of S_p is S_{p-1}. The set $S_p \setminus S_{p-1}$ of states of rank equal to p is then a locally closed subvariety.

There is also a nice analytic characterization of pure states ϕ in $H_1 \otimes H_2$, in terms of the projection operator P_ϕ. The partial trace $\rho := Tr_1(P_\phi)$ is a positive hermitian operator on H_2 and we have:

Proposition 1.5. *We have $\rho^2 \leq \rho$ with equality iff ϕ is pure.*

For the proof see Popescu-Rohrlich [Po-Ro]. There is an interesting relation with the algebraic characterization of pure states in Proposition 1.1, which we illustrate for $\mathbb{C}^2 \otimes \mathbb{C}^2$,

using the basis (e_0, e_1) of $\mathbb{C}^2$. Here ϕ is given by a matrix (ϕ_{ab}). The $(0,0)$-component of $\rho - \rho^2$ is equal to $|\phi_{00}\phi_{11} - \phi_{01}\phi_{10}|^2$. Thus the analytic equations characterizing pure states are quartic real polynomials which are squares (in general, sums of squares) of absolute values of the quadratic complex polynomial equations.

For $k > 2$ the situation is more complicated: it is always true that $S_p \subset S_{p+1}$, but we will see in the next section that the set S_p is not always closed in $\mathbb{P}(E)$.

Note that for $k = 3$ the rank of a tensor is closely connected to the notion of rank of a bilinear map [Str1] [Str2] [B-C-S].

2. Tensors in $(\mathbb{C}^2)^{\otimes 3}$.

We study here $E = \mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^2$. There is a well-known polynomial function D on E which is invariant under $SL(2,\mathbb{C})^3$: this is the hyperdeterminant introduced by Cayley [Cay] [G-K-Z]. D is a homogeneous degree 4 polynomial function on E which is $SL(2,\mathbb{C})^3$-invariant. We pick the standard basis e_0, e_1 of $\mathbb{C}^2$ and write the components of $u \in (\mathbb{C}^2)^{\otimes 3}$ as u_{abc} for $a, b, c \in \{0, 1\}$. Then we have:

$$\begin{aligned} D(u) = \ & u_{000}^2 u_{111}^2 + u_{001}^2 u_{110}^2 + u_{010}^2 u_{101}^2 + u_{011}^2 u_{100}^2 \\ & -2(u_{000}u_{001}u_{110}u_{111} + u_{000}u_{010}u_{101}u_{111} + u_{000}u_{011}u_{100}u_{111} \\ & +u_{001}u_{010}u_{101}u_{110} + u_{001}u_{011}u_{110}u_{100} + u_{010}u_{011}u_{101}u_{100}) \\ & +4(u_{000}u_{011}u_{101}u_{110} + u_{001}u_{010}u_{100}u_{111}) \end{aligned} \tag{2.1}$$

The geometric significance of D is that $D(u) = 0$ iff the hyperplane defined by u is tangent to the Segre product S at some point p. This means that $\langle u|v\rangle = 0$ for any tangent vector v to S at p.

The review [Cat] provides interesting comments on the book [G-K-Z].

For a tensor u in $E = (\mathbb{C}^2)^{\otimes 3}$, there are three additional degrees of entanglement $\delta_1, \delta_2, \delta_3$ to consider: δ_1 is the rank of u viewed as an element of $\mathbb{C}^2 \otimes \mathbb{C}^4$, when we group the second and third factors $\mathbb{C}^2$. δ_2 and δ_3 are defined similarly.

We denote by Y_j the closed algebraic subvariety of $\mathbb{P}(E)$ comprised of states u such that $\delta_j = 1$, i.e., u belongs to Y_1 iff it is decomposable as a tensor in $\mathbb{C}^2 \otimes \mathbb{C}^4$. Note each Y_j has dimension 4 and is contained in the hypersurface Z of equation $D = 0$.

The following result is proved (at least implicitly) in [G-K-Z]. We include a proof since it uses methods which we will later use for $(\mathbb{C}^2)^{\otimes 4}$.

Proposition 2.1. *Let u be a state in $E = (\mathbb{C}^2)^{\otimes 3}$. Then u satisfies exactly one the following possibilities:*

(1) u is a pure state.

(2) u is not pure but belongs to Y_j for a (unique) $j = 1, 2, 3$.

(3) u is entangled, and $D(u) \neq 0$; in that case u has rank 2, so it is the sum of two pure tensors

(4) $D(u) = 0$, but u belongs to none of the Y_j; then u has rank 3.

Proof. It is useful to associate to u a linear map $T : \mathbb{C}^2 \to \mathbb{C}^2 \otimes \mathbb{C}^2$. If T has rank 1 then $u \in Y_1$ and we are in case (1). So we may assume T has rank 2. We will consider $T(xe_0 + ye_1)$ as a 2 by 2 matrix. Consider the homogeneous degree 2 polynomial $P(x, y) = det(T(xe_0 + ye_1))$. There are 3 cases to consider:

(I) there are exactly two points in $\mathbb{CP}^1$ where P vanishes. Let $(x_1, y_1), (x_2, y_2)$ be homogeneous coordinates for these two points. Then we can make a change a basis in the first $\mathbb{C}^2$ so that these 2 points are $(0, 1)$ and $(1, 0)$. Then both $T(e_0)$ and $T(e_1)$ have rank ≤ 1; they must both be non-zero, otherwise all $T(xe_0 + ye_1)$ would have rank ≤ 1. After a change of basis in the second and third copies of $\mathbb{C}^2$, we may assume $T(e_0) = e_0 \otimes e_0$ and $T(e_1) = e_i \otimes e_j$ for suitable i, j, not both equal to 0. In this case, the tensor u is equal to $e_0 \otimes e_0 \otimes e_0 + e_1 \otimes e_i \otimes e_j$, so it has rank equal to 2. By direct computation, we see that $D(u) \neq 0$ if $i = j = 1$ (case 4), or u belongs to Y_2 (resp. Y_3) if $i = 0$ (resp. $j = 0$), which belongs to case (2)..

(II) there is only one point (x, y) of $\mathbb{CP}^1$ at which $P(x, y)$ vanishes. We may assume this point is $(1, 0)$. We can think of T as giving a parameterization of a curve in $\mathbb{CP}^3$ which is tangent to the quadric surface Q consisting of rank 1 matrices. We can change bases in all three copies of $\mathbb{C}^2$ so that $T(e_0) = e_0 \otimes e_0$. As the tangent plane to Q at $e_0 \otimes e_0$ is spanned by $e_0 \otimes e_1$ and $e_1 \otimes e_0$, we can change the basis vector e_1 in the first $\mathbb{C}^2$ so that $T(xe_0 + ye_1) = xe_0 \otimes e_0 + y(\lambda e_0 \otimes e_1 + \mu e_1 \otimes e_0)$. Next, as λ and μ must both be non-zero, we can change bases in the other copies to arrange that $\lambda = \mu = 1$. Then our tensor u is $u = e_0 \otimes e_0 \otimes e_0 + e_1 \otimes e_1 \otimes e_0 + e_1 \otimes e_0 \otimes e_1$, and has rank exactly 3. Indeed it has the property that for any non-zero $v \in \mathbb{C}^2$, the tensor in $\mathbb{C}^2 \otimes \mathbb{C}^2$ obtained by contracting u with w has rank equal to 2; thus u can't be a sum of two pure tensors. We verify easily that $D(u) = 0$. Or we can see geometrically that the corresponding point in $\mathbb{P}(E)$ belongs to the dual variety to the Segre product $S = S_1$, which means that the hyperplane defined by u is tangent to the Segre variety at some point. The relevant point of S is $v = e_0 \otimes e_1 \otimes e_1$: notice that the tangent space to S at v is spanned by tensors of the type $\psi \otimes e_1 \otimes e_1, e_0 \otimes \psi \otimes e_1, e_0 \otimes e_1 \otimes \psi$ where $\psi \in \mathbb{C}^2$. Since u is orthogonal to all these tangent vectors, it is orthogonal to the tangent space of S at v.

(III) the polynomial $P(x, y)$ vanishes identically; this means that the linear map $T(xe_0 + ye_1)$ always has rank ≤ 1. This can happen in either of 2 ways:

(a) there is a vector $\psi \in \mathbb{C}^2$ and a linear map $f : \mathbb{C}^2 \to \mathbb{C}^2$ such that $T(w) = \psi \otimes f(w)$

(b) there is a vector $\psi \in \mathbb{C}^2$ and a linear map $f : \mathbb{C}^2 \to \mathbb{C}^2$ such that $T(w) = f(w) \otimes \psi$

We need only consider case (a). Then we have $u = e_0 \otimes \psi \otimes f(e_0) + e_1 \otimes \psi \otimes f(e_1)$. If $f(e_0)$ and $f(e_1)$ are linearly dependent, the tensor u is pure and we are in case (1). Otherwise, u has rank 2 and after changes of bases in the second and third copies of $\mathbb{C}^2$ it takes the form $u = e_0 \otimes e_0 \otimes e_0 + e_1 \otimes e_0 \otimes e_1$. Then u belongs to Y_2.

□

This also leads to normal forms for tensor states in $(\mathbb{C}^2)^{\otimes 3}$ up to the action of $GL(2, \mathbb{C})^3$; these normal forms are given in [G-K-Z]. For quantum mechanics one needs to consider the smaller symmetry group of unitary symmetries $U(2)^3$. This is the group of local unitary symmetries; we say that two tensor states are *locally equivalent* if they are equivalent under $U(2)^3$. One obtains normal expressions up to local equivalence:

Proposition 2.2. *A state in $(\mathbb{C}^2)^{\otimes 3}$ is locally equivalent to one of the following:*

1) a pure state is locally equivalent to $e_0 \otimes e_0 \otimes e_0$.

2) a state in Y_1 which is not pure is locally equivalent to

$$e_0 \otimes [\cos\theta \ (e_0 \otimes e_0) + \sin\theta \ (e_1 \otimes e_1)] \tag{2.2}$$

States in Y_2 or Y_3 are described similarly.

3) a state of rank 2 which is not in either of the Y_j is locally equivalent to

$$\lambda \ e_0 \otimes e_0 \otimes e_0 + z(\cos\theta_1 e_0 + \sin\theta_1 e_1) \otimes (\cos\theta_2 e_0 + \sin\theta_2 e_1) \otimes (\cos\theta_3 e_0 + \sin\theta_3 e_1) \tag{2.3}$$

where $\lambda, \theta_j \in \mathbb{R}$, $z \in \mathbb{C}$ satisfy the relation $\lambda^2 + |z|^2 + 2\lambda\Re(z)\cos\theta_1\cos\theta_2\cos\theta_3 = 1$ (so that the tensor has norm 1). We can assume $\theta_j \in (0, \frac{\pi}{2})$.

4) a state of rank 3 is locally equivalent to

$$\cos\theta_1 e_0 \otimes e_0 \otimes e_0 + \sin\theta_1 e_1 \otimes [\cos\theta_2 e_0 \otimes (\cos\theta_3 e_0 + \sin\theta_3 e_1) + \sin\theta_2 e_1 \otimes e_0] \tag{2.4}$$

In each case there are only finitely many values of the parameters corresponding a given tensor state.

Proof. The four cases of the statement correspond to the four cases of Proposition 2.1. Case 1) is obvious. Case 2) follows as $u \in Y_1$ is locally equivalent to $e_0 \otimes v$ for some $v \in \mathbb{C}^2 \otimes \mathbb{C}^2$; by Schmidt's theorem v is locally equivalent to $\cos\alpha \ (e_0 \otimes e_0) + \sin\alpha \ (e_1 \otimes e_1)$.

In case 3), we have $u = v_1 \otimes v_2 \otimes v_3 + w_1 \otimes w_2 \otimes w_3$ where v_i and w_i are linearly independent for each i. There is no harm in assuming that the vectors v_1, v_2, w_1, w_2 have norm 1. By rescaling u by a phase and applying a local transformation we can assume $v_1 = v_2 = e_0$ and $v_3 = \lambda e_0$ for $\lambda > 0$. We can also arrange that $w_i = \cos\theta_i e_0 + \sin\theta_i e_1$ for $i = 1, 2$ and $w_3 = z[\cos\theta_3 e_0 + \sin\theta_3 e_1]$ for some $z \in \mathbb{C}$. This gives the normal form; note the reduction to $\theta_j \in (0, \frac{\pi}{2})$ is easy to achieve by changing the signs of e_0 and e_1 in the j-th factor $\mathbb{C}^2$.

In case 4), the tensor u has the form $u = v_1 \otimes v_2 \otimes v_3 + w_1 \otimes (v_2 \otimes w_3 + w_2 \otimes v_3)$ where v_i and w_i are linearly independent for each i. There are two types of degrees of freedom in the expression of u in this form. First we have the transformation $u = (v_1 + \alpha w_1) \otimes v_2 \otimes v_3 + w_1 \otimes (v_2 \otimes [w_3 - \alpha v_3] + w_2 \otimes v_3)$. With its help we can arrange that $w_1 \perp v_1$. Secondly we have $u = v_1 \otimes v_2 \otimes v_3 + w_1 \otimes (v_2 \otimes [w_3 + \beta v_3] + [w_2 - \beta v_2] \otimes v_3)$. This is used to arrange that $w_2 \perp v_2$. By rescaling the v_i, we may assume that v_1 and v_2 have norm 1. After applying a local unitary transformation, we obtain $v_1 = v_2 = e_0$ and $v_3 = \lambda e_0$ for $\lambda \in \mathbb{C}^*$. Then we have $w_1 = \alpha e_1$ and $w_2 = \beta e_1$ for suitable $\alpha, \beta \in \mathbb{C}^*$. Write $w_3 = \mu(\cos\theta e_0 + \sin\theta e_1)$ for $\mu \in \mathbb{C}^*$. Thus we have $u = \lambda e_0^{\otimes 3} + \alpha e_1 \otimes (\mu e_0 \otimes [\cos\theta e_0 + \sin\theta e_1] + \nu e_1 \otimes e_0)$ for some $\nu \in \mathbb{C}$. Clearly a phase change for u will make λ real, so we can assume $\lambda \in \mathbb{R}$. We can of course assume $\alpha = 1$ by changing μ and ν appropriately.

In the rest of the proof we use the notation $e_i^{(j)}$ to denote the vector e_i in the j-th copy of $\mathbb{C}^2$. We will next do a simultaneous phase change

$$e_0^{(1)} \mapsto e^{i\phi} e_0^{(1)}, e_0^{(3)} \mapsto e^{-i\phi} e_0^{(3)}, e_1^{(3)} \mapsto e^{-i\phi} e_1^{(3)} \tag{2.5}$$

This operation does not change λ but rescales μ as well as ν; so we can assume μ is real. Finally a phase rescaling of $e_1^{(2)}$ will make ν real without changing λ or μ. This way we easily get the normal form.

□

It is interesting to discuss why S_2 is not closed in the case of $(\mathbb{C}^2)^{\otimes 3}$. There is an easy geometric description of the rank, which is well-known to algebraic geometers. We start with the Segre product $S = S_1$, which is a closed algebraic subvariety of $\mathbb{P}(E)$. For a $(p-1)$-plane $\Pi \subset \mathbb{P}(E)$, we say that Π is a p-secant plane if Π is spanned by p points $y_1, \cdots, y_p$ of $\Pi \cap S$. For instance, a line is 2-secant if it is a secant line, a 2-plane is 3-secant if it spanned by 3 points of $\Pi \cap S$. Then we have clearly

Lemma 2.3. *A point of $\mathbb{P}(E)$ has rank $\leq p$ iff it belongs to some $(p-1)$-plane $\Pi \subset \mathbb{P}(E)$ which is p-secant to the Segre product S. In other words S_p is the union of all $(p-1)$-planes Π which are p-secant.*

The point then is that S_2 need not be closed, because the limit of a sequence of secant lines to S need not be a secant line, but could be a tangent line. This is similar to the fact that the border rank of a bilinear map can be lower than its rank [B-C-L-R] [Str2]. The same phenomenon could occur for higher p. From algebraic geometry we have the following general fact. In this statement, we use the Zariski topology of $\mathbb{P}(E)$ for which the closed subsets are the subsets defined by homogeneous polynomial equations. The constructible sets are then those obtained from the Zariski closed subsets by finite Boolean operations (finite unions, finite intersections, and complementation). A closed subset F is called irreducible if whenever $F = G \cup H$ for G, H closed, we have $G = F$ or $H = F$.

Another interesting phenomenon is that a real tensor in $(\mathbb{R}^2)^{\otimes 3}$ may have different rank from the same tensor viewed as an element of $(\mathbb{C}^2)^{\otimes 3}$. An example is the tensor $e_0 \otimes (e_0 \otimes e_0 - e_1 \otimes e_1) + e_1(e_0 \otimes e_1 + e_1 \otimes e_0)$, which corresponds to the product law $\mathbb{R}^2 \otimes \mathbb{R}^2 \to \mathbb{R}^2$ on $\mathbb{R}^2 = \mathbb{C}$; this has rank 3 as a real tensor and rank 2 as a complex tensor.

Proposition 2.4. *The set S_p is a Zariski constructible subset of $\mathbb{P}(E)$. The closure $\bar{S}_p$ is irreducible.*

Proof. Let T_p be the image of an algebraic mapping $\Phi : X \to \mathbb{P}(E)$, where $X \subset S^p \times \mathbb{P}(E)$ is the locally closed algebraic subvariety comprised of $p+1$-uples $(x_1, \cdots, x_{p+1})$ where $x_1, \cdots, x_p \in S$ are distinct, $x_{p+1} \in \mathbb{P}(E)$ and $(x_1, \cdots, x_{p+1})$ belong to some $(p-1)$-plane, and $\Phi(x_1, \cdots, x_{p+1}) = x_{p+1}$. It is easy to see that X is irreducible; thus standard results in algebraic geometry say that T_p is constructible and its closure is irreducible. We have easily $S_p = \cup_{q \leq p} T_q$ so that S_p is constructible. It is clear that $\bar{T}_q \subseteq \bar{T}_p$ for $q \leq p$, so that $\bar{S}_p = \bar{S}_p$ is irreducible.

□

The method of proof of Proposition 2.1 leads naturally to the following notion

Definition 2.5. Let F be subspace of $E = H_1 \otimes \cdots \otimes H_k$. The rank of F is the smallest integer p such that there exist p pure tensors $u_1, \cdots, u_p$ such that F is contained in the span of $u_1, \cdots, u_p$.

We then have the following easy but useful result:

Lemma 2.6. [B-C-S, Prop. 14.44] *Let $u \in E = H_1 \otimes H_2 \otimes \cdots \otimes H_k$, and let T be the corresponding linear map $T : H_1^* \to H_2 \otimes \cdots \otimes H_k$. Then the rank of the tensor u is equal to the rank of the range of T as a subspace of $H_2 \otimes \cdots \otimes H_k$.*

Proof. Let p be the rank of u and q the rank of the range of T. Thus u is a linear combination of pure tensors $v_1, \cdots, v_p$. Write $v_j = w_j \otimes z_j$ where $w_1 \in H_1$ and $z_j \in H_2 \otimes \cdots \otimes H_k$. Then we have $T(l) = \sum_j \langle l|w_j\rangle z_j$, so that $T(l)$ is a linear combination of the pure tensors z_j and $q \leq p$. In the other direction, assume that the range of T is contained in the linear span of the pure tensors $\beta_j, 1 \leq l \leq s$. Then there are linear forms v_j on H_1^* (so $v_j \in H_1$) such that $T(l) = \sum_{j=1}^s \langle l|v_j\rangle\beta_j$. This means that $u = \sum_{j=1}^s v_j \otimes \beta_j$ and $r \leq s$.

□

There is a classical example for the rank of a subspace of $M_2(\mathbb{C})^{\otimes 2}$. We identify $M_2(\mathbb{C})$ with its dual, so that $M_2(\mathbb{C})^{\otimes 2}$ identifies with the space of bilinear functionals $(A, B) \mapsto f(A, B)$ of two matrices A, B of size 2. The coefficients of the product AB yield four such bilinear functionals, which span a four-dimensional subspace E of $M_2(\mathbb{C})^{\otimes 2}$. It is a well-known result of Strassen [Str1] [Str2] that the rank of this subspace is equal to 7 (instead of the value 8 one might naively expect). This is the basis for fast matrix multiplication. From Lemma 2.6 it ensues that the corresponding tensor in $M_2(\mathbb{C}) \otimes M_2(\mathbb{C}) \otimes M_2(\mathbb{C}) = M_2(\mathbb{C})^{\otimes 3}$ has rank equal to 7. It would be nice to have a geometric interpretation of this fact.

We also note an easy consequence of Lemma 2.6

Lemma 2.7. *Let $(e_1, \cdots e_{d_1})$ be a basis of H_1, and consider a tensor $u = \sum_j e_j \otimes v_j \in H_1 \otimes H_2 \otimes \cdots \otimes H_k$. Then the rank of u is at most the sum of the ranks of the v_j's.*

3. Tensors in $(\mathbb{C}^2)^{\otimes 4}$.

Our results for $E = (\mathbb{C}^2)^{\otimes 4}$ are less complete than for $(\mathbb{C}^2)^{\otimes 3}$. For $E = (\mathbb{C}^2)^{\otimes 4}$, an important invariant is the following: for any permutation (i, j, k, l) of $(1, 2, 3, 4)$, a tensor $u \in (\mathbb{C}^2)^{\otimes 4}$ yields a linear map $\phi_{ijkl} : \mathbb{C}^2 \otimes \mathbb{C}^2 \to \mathbb{C}^2 \otimes \mathbb{C}^2$. We consider the determinant $\Delta(ijkl) = det(\phi_{ijkl})$. We have the following symmetries: $\Delta(ijkl) = -\Delta(jikl) = -\Delta(ijlk) = \Delta(klij)$, so up to sign we have essentially 3 determinants. Now in fact we have

Lemma 3.1.

$$\Delta(1234) - \Delta(1324) + \Delta(1423) = 0. \tag{3.1}$$

Proposition 3.2. *Let $E = (C^2)^4$. Then the closure of S_3 in $\mathbb{P}(E)$ is the dimension* 13 *algebraic variety defined by the equations $\Delta(ijkl) = 0$*

Proof. The Lie group $G = GL(2, \mathbb{C})$ acts naturally on E and preserves each S_p. Let T be the subspace spanned by $e_0^{\otimes 4}$, $e_1^{\otimes 4}$ and $(e_0 + e_1)^{\otimes 4}$.

Clearly the closure of S_3 is the closure of the G-saturation $G \cdot T$. We can compute its dimension as follows. We consider the infinitesimal equation of the Lie algebra $\mathfrak{g} = \mathfrak{gl}(4, \mathbb{C})$ on E. For $v \in T$, we denote by $\mathfrak{h}_v$ the space of $\gamma \in \mathfrak{g}$ such that $\gamma \cdot v \in T$. Then we have

$$\dim(G \cdot T) = \dim(G) + \dim(S) - min_{v \in S} \dim(\mathfrak{h}_v) - 1 = 18 - min_{v \in S} \dim(\mathfrak{h}_v) \tag{3.2}$$

This follows as the right-hand side is the rank of the mapping $G \times T \to T \to \mathbb{P}(E)$ at the point $(1, v)$.

Now for any $\delta, \epsilon \in \mathbb{C}^*$, the tensor $u = u_{\delta,\epsilon} = e_0^{\otimes 4} + \delta e_1^{\otimes 4} + \epsilon(e_0 + e_1)^{\otimes 4}$ belongs to T. Let $\mathfrak{k}_u$ be the space comprised of the $\gamma \in \mathfrak{g}$ such that $\gamma \cdot u$ is a linear combination of $e_0^{\otimes 4}$ and $e_1^{\otimes 4}$. Since $\mathfrak{h}_v$ is the direct sum of $\mathfrak{k}_v$ and of the line spanned by $(Id, 0, 0, 0)$, we have $\dim(\mathfrak{h}_u) = \dim(\mathfrak{k}_u) + 1$. So it suffices to compute $\dim(\mathfrak{k}_u)$.

Now for $\gamma = (\gamma_j) \in \mathfrak{g} = \mathfrak{gl}(2, \mathbb{C})^4$ with $\gamma_j = \begin{pmatrix} a_j & b_j \\ c_j & d_j \end{pmatrix}$ we compute:

$$\begin{aligned} \gamma \cdot u \; &= (a_1 + a_2 + a_3 + a_4 + \epsilon(a_1 + a_2 + a_3 + a_4 + b_1 + b_2 + b_3 + b_4))\ e_0^{\otimes 4} \\ &+ (c_4 + \epsilon(a_1 + b_1 + a_2 + b_2 + a_3 + b_3 + c_4 + d_4))\ e_0^{\otimes 3} \otimes e_1 + \text{permutations} \\ &+ (\delta d_1 + \epsilon(a_1 + b_1 + c_2 + d_2 + c_3 + d_3 + c_4 + d_4)) e_0 \otimes e_1^{\otimes 3} + \text{permutations} \\ &+ \epsilon(a_1 + b_1 + c_2 + d_2 + c_3 + d_3 + c_4 + d_4) e_0^{\otimes 2} \otimes e_1^{\otimes 2} + \text{permutations} \end{aligned} \tag{3.3}$$

So γ belongs to $\mathfrak{k}_v$ iff the coefficients of $e_0^{\otimes 3} \otimes e_1$, $e_0^{\otimes 2} \otimes e_1^{\otimes 2}$, $e_0 \otimes e_1^{\otimes 3}$, and the other tensors obtained from these by permutations, all vanish. At first sight this is just a system of linear equations in 16 unknowns, but one can essentially separate them according the four groups of four variables (a_j, b_j, c_j, d_j) by introducing the sums $\lambda = \sum_j a_j, \mu = \sum_j b_j, \nu = \sum_j c_j, \rho = \sum_j d_j$. One gets the equations:

(1) For each j, $\epsilon(a_j + b_j - d_j) - (1 + \epsilon)c_j = \epsilon(\lambda + \mu)$
(2) For each j, $\epsilon(a_j - c_j - d_j) + (\delta + \epsilon)b_j = \epsilon(\nu + \rho)$
(3) for each permutation $(ijkl)$ of (1234) we have $a_i + b_i + a_j + b_j + c_k + d_k + c_l + d_l = 0$.
(3) easily implies that $a_j + b_j - c_j - d_j$ is independent of j.

By summing each of the three types of equations over all choices of indices (or of permutations for the third), we get consistency requirements for $(\lambda, \mu, \nu, \rho)$; these are easily solved to yield:

$$\mu = \frac{2\epsilon}{\delta - 2\epsilon}\lambda, \nu = \frac{-2\delta\epsilon}{\delta - 2\epsilon}\lambda, \rho = \frac{\lambda\delta(2\epsilon - 1)}{\delta - 2\epsilon}, \tag{3.4}$$

Here λ is a free parameter; once it is chosen we can solve for (a_j, b_j, c_j, d_j) and obtain $\gamma_j = \omega_j Id + \phi_j \xi + \eta$, where $\xi = \begin{pmatrix} \delta + \epsilon + \delta\epsilon & -\epsilon \\ \delta\epsilon & 0 \end{pmatrix}$, $\eta = \begin{pmatrix} \frac{\delta\lambda}{\delta - 2\epsilon} & 0 \\ 0 & 0 \end{pmatrix}$ are matrices independent of j, and ω_j, ϕ_j are some scalars. The fact that $a_j + b_j - c_j - d_j$ is independent of j then implies that ϕ_j is too; call this scalar ϕ. Then we need the a_j to sum up to λ, etc... This gives the value $\frac{-\mu}{4\epsilon}$ for ϕ and the requirement $\sum \omega_j = \rho$.

Counting the free parameters we obtain $\dim(\mathfrak{k}_u) = 4$. It follows that S_3 has dimension 13. It is clearly contained in the codimension 2 subvariety defined by the vanishing of the $\Delta(ijkl)$; the latter variety is seen to be irreducible, thus it must equal the closure of S_3. □

Theorem 3.3. *The highest rank of a tensor in $(\mathbb{C}^2)^{\otimes 4}$ is equal to 4.*

Proof. We associate to $u \in (\mathbb{C}^2)^{\otimes 4}$ as before a linear map $T : \mathbb{C}^2 \to (\mathbb{C}^2)^{\otimes 3}$ and compute the rank of its image. If T has rank 1 it is clear that u has rank ≤ 3, so we can assume T is injective. We can think of ϕ as parameterizing a line l in $\mathbb{P}((\mathbb{C}^2)^{\otimes 3}) = \mathbb{CP}^7$. If this line is not contained in the hypersurface Z, then 2 of its points have rank ≤ 2, and it follows using Lemma 2.7 that u has rank $\leq 2 + 2 = 4$. Thus we need to focus on the case where l is contained in Z. First of all, there is the case where l is contained in

Y_j for some j. In that case it is easy to see that u is of rank ≤ 4. So we can assume that l contains a point v which belongs to none of the Y_j; so v is $GL(2,\mathbb{C})^3$-conjugate to $e_0 \otimes e_0 \otimes e_0 + e_1 \otimes e_0 \otimes e_1 + e_1 \otimes e_1 \otimes e_0$. For this choice of vector v we can write down the equations on a tensor w so that the line thru v and w is contained in Z, i.e., $D(xv + yw)$ vanishes identically. It is natural to consider w as a vector modulo scaling in the quotient space $(\mathbb{C}^2)^{\otimes 3}/\mathbb{C}v = \mathbb{C}^7$, i.e., as an element of projective space $\mathbb{CP}^6$. Look at the equation giving the vanishing of the coefficient of $x^i y^{4-i}$, as i ranges from 3 to 0. The first equation is $w_{011} = 0$. The second is $2(w_{111}^2 + w_{001}^2 + w_{010}^2) - (w_{001} + w_{010} + w_{111})^2 = 0$; this is a non-degenerate quadratic form in 3 variables. The third equation involves the new variables $w_{000}, w_{101}, w_{110}$ and is linear as a function of these 3 variables. The fourth equation involves also the last variable w_{100}, and is linear in w_{100}. The variety of w such that $D(xv + yw) \equiv 0$ thus has a dense open set which is obtained by successive fibrations with fibers irreducible algebraic varieties; thus it itself is irreducible and its dimension is equal to 2. Denote by Z^0 the big $SL(2,\mathbb{C})^3$-orbit inside Z, which is the complement of $Y_1 \cup Y_2 \cup Y_3$. Now consider the algebraic variety S comprised of pairs (p, L) where $p \in Z^0$ and L is a line thru p which lies entirely inside Z. This is a locally closed subvariety of the product of Z^0 with the Grassmann manifold of lines in $\mathbb{CP}^7$. Then the projection map $S \to Z^0$ is a fibration, because it is $SL(2,\mathbb{C})^3$-equivariant and Z^0 is a single orbit. The dimension of S is therefore $6 + 2 = 8$. What we are really after however is the variety V of lines contained in Z and meeting Z^0. There is an obvious map $S \to V$ which is a smooth mapping with one-dimensional fibers. Therefore V has dimension $8 - 1 = 7$. Now we claim that any line contained in Z and not contained in any Y_j must meet each Y_j. For this purpose consider some tensor in Y_1, say $v = e_0^3 + e_0 \otimes e_1^2$, and consider again the set of w such that $D(xv + yw) \equiv 0$.

One checks that this forms a subvariety of $\mathbb{P}^6$ of dimension 3. It follows that the set of lines contained in Z and meeting Y_1 in finitely many points has a finite ramified covering which maps to Y_1 with three-dimensional fiber, therefore it has dimension $4 + 3 = 7$. Note that the lines completely contained in Y_1 form a variety of dimension 5. It then follows that any line contained in Z must meet each Y_j.

Thus we can change the basis of the first $\mathbb{C}^2$ so that $T(e_0) \in Y_1$ and $T(e_1) \in Y_2$. Then both these tensors have rank ≤ 2, and by Lemma 2.7 u itself has rank $\leq 2 + 2 = 4$.

□

It is easy to see that S_2 has dimension 9 and satisfies a number of algebraic equations, namely the 2 by 2 minors of the linear maps $(\mathbb{C}^2)^{\otimes 2} \to (\mathbb{C}^2)^{\otimes 2}$ obtained from the tensor (there are essentially 3 such linear maps). It is likely the case that these equations precisely describe the closure of S_2.

References

[B-C-L-R] Bini, Capovanni, Lotti and Romani, $O(n^{2.7799})$ complexity for matrix multiplication, Inf. Proc. Letters 8 (1979), 234-235

[B-C-S] P. Bürgisser, M. Clausen and M. A. Shokrollahi, Algebraic Complexity Theory, Grundl. vol. 315, Springer-verlag (1997)

[C-R-S-S] A. R. Calderbank, E. M Rains, P. W. Shor, N. J. A. Sloane, Quantum Error Correction via Codes over GF(4), IEEE Trans. Inform. Theory 44 (1998), 1369-1387; quant-ph/9608006

[Cat] F. Catanese, Review of the book [C-R-S-S], Bull. Amer. Soc. (2000)

[Cay] A. Cayley, On the theory of elimination, Collected Papers vol. 1, no. 59, Cambridge Univ. Press (1889), 370-374

[G-K-Z] I.M. Gelfand, M. Kapranov and A. Zelevinsky, Discriminants, Resultants and Multidimensional Deterrminants, Birkhäuser (1991)

[Go] D. Gottesman, An Introduction to Quantum Error Correction, talk given at AMS Short Course on Quantum Computation in Jan. 2000, quant-ph/0004072

[Po-Ro] S. Popescu and R. Rohrlich, The joy of entanglement, Introduction to Quantum Computation and Information, H-K Lo, S. Popescu, T. Spiller eds, World Scientific (1998), 29-48

[Ste] A. Steane, Simple Quantum Error Correcting Codes, Phys.Rev. A54 (1996), 4741

[Str1] V. Strassen, Gaussian elimination is not optimal, Numer. Math. 13 (1969), 354-356

[Str2] V. Strassen, Rank and optimal computation of general tensors, Linear Alg. Appl. 52/53 (1983), 645-685

Department of Mathematics, Penn State University, University Park 16802

E-mail address: `jlb@math.psu.edu`

URL: `www.math.psu.edu/jlb`